

Title: Intersentential coreference expectations reflect mental models of events

Article type: Short communication

Authors: Theres Grüter<sup>a</sup> (theres@hawaii.edu; corresponding author)  
Aya Takeda<sup>a</sup> (ayataked@hawaii.edu)  
Hannah Rohde<sup>b</sup> (hannah.rohde@ed.ac.uk)  
Amy J. Schafer<sup>a</sup> (amy.schafer@hawaii.edu)

Affiliations: <sup>a</sup> University of Hawai'i at Mānoa

<sup>b</sup> University of Edinburgh

Address for correspondence:

Theres Grüter  
Department of Second Language Studies  
University of Hawai'i at Mānoa  
1890 East-West Road, Moore Hall, rm570  
Honolulu, HI 96822  
U.S.A.

Word count: 2,948  
(including abstract, main text, and figure captions; excluding title, author affiliations, keywords, acknowledgments, footnotes, bibliography and figures)

Accepted for publication in *Cognition*, April 4, 2018

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Abstract: Comprehenders' perception of the world is mediated by the mental models they construct. During discourse processing, incoming information allows comprehenders to update their model of the events being described. At the same time, comprehenders use these models to generate expectations about who or what will be mentioned next. The temporal dynamics of this interdependence between language processing and mental event representation has been difficult to disentangle. The present visual world eye-tracking experiment measures listeners' coreference expectations during an intersentential pause between a sentence about a transfer-of-possession event and a continuation mentioning either its Source or Goal. We found a temporally dispersed but sustained preference for fixating the Goal that was significantly greater when the event was described as completed rather than incomplete (*passed* versus *was passing*). This aligns with reported offline sensitivity to event structure, as conveyed via verb aspect, and provides new evidence that our mental model of an event leads to early and, crucially, proactive expectations about subsequent mention in the upcoming discourse.

Keywords: coreference; expectation-driven processing; verb aspect; eye tracking; mental model; event representation

## 1. Introduction

When we process discourse, we create mental models of the events being described (Johnson-Laird, 1983; Van Dijk & Kintsch, 1983). We use incoming linguistic information, together with our knowledge of the world, to incrementally update these models as the discourse progresses. Subtle differences in linguistic choices can have measurable effects on our mental representations of a situation. For example, events described with imperfective-marked verbs (e.g., *Leah was passing the salt to Eve*) are construed as ongoing, with attention distributed across event participants, whereas perfective aspect (*Leah passed the salt to Eve*) evokes an event as completed, bringing participants associated with its end state into the focus of attention (Madden & Ferretti, 2009; Magliano & Schleich, 2000; Moens & Steedman, 1988). In the domain of reference processing, this focus of attention on an event participant is linked to the probability of that participant being rementioned in the next sentence. An ambiguous pronoun following a perfective-marked transfer-of-possession event (e.g., *Leah passed the salt to Eve. She...*) is preferentially interpreted as referring to the Goal of the transfer event (*Eve*), who is now in possession of the transferred object, rather than the Source (*Leah*; Stevenson, Crawley, & Kleinman, 1994). Importantly, when the verb is marked with imperfective (*was passing*), this Goal preference decreases (Rohde, Kehler, & Elman, 2006; Kehler, Kertz, Rohde, & Elman, 2008), consistent with more equal focus on participants in a mental model of an *ongoing* event.

In this paper we investigate this interdependence between language processing and our mental representations of events, probing the timecourse over which our situation models inform our processing of reference. This question has been addressed extensively

for events and situations described with Implicit Causality verbs (IC; Garvey & Caramazza, 1974; Hartshorne, 2014), where a long-standing debate centers on whether these referential biases emerge before, at, or after a pronoun in contexts involving inferences about the event's cause (see Koornneef, Dotlačil, van den Broek, & Sanders, 2016, for review).<sup>1</sup> One possibility is that these biases depend on the referential expression, such that a pronoun prompts comprehenders to consult their current situation model, and the distribution of attention to event participants in that model influences their referential choices. This could explain findings from story continuation experiments in which manipulations of IC status or grammatical aspect influenced participants' referential choices when they wrote completions for sentences starting with an ambiguous pronoun (Kehler et al., 2008). It is also compatible with findings from an ERP study by Ferretti, Rohde, Kehler and Crutchley (2009), in which participants read sentences containing perfective- or imperfective-marked transfer-of-possession verbs with protagonists of different gender (e.g., *Sue handed/was handing the timecard to Fred*), and a subsequent sentence-initial pronoun (*He/She*) immediately disambiguated reference. Ferretti and colleagues reported an enhanced P600 at the pronoun when it matched the gender of the Source (*Sue*) following perfective but not imperfective events.

---

<sup>1</sup> IC verbs are interpersonal predicates whose meaning and use crucially involve a primary attribution of cause to one event participant. The syntactic position of the causally-implicated referent distinguishes subject-biased IC verbs (*Max frightened Ben because he...* [bias to Max]) and object-biased IC verbs (*Max feared Ben because he...* [bias to Ben]; Ferstl, Garnham, & Manouilidou, 2011), at least with an explicit *because* (cf. Koornneef & Sanders 2013).

This indicates comprehenders had more difficulty integrating the pronoun when its reference (forced by gender-marking) jarred with the status of that referent in their situation model (for similar effects on reading time in contexts with IC verbs, see Koornneef & van Berkum, 2006, and Koornneef & Sanders, 2013).

Another account of these findings, and the one advocated by Ferretti, Rohde and colleagues, postulates a proactive bias whereby comprehenders' situation models exert an influence on referential processing independent of the presence of a particular referential form. More specifically, comprehenders may continually draw on their current situation models to generate expectations about who or what is likely to be mentioned next in the upcoming discourse. If and when a pronoun is encountered, its interpretation is then in part a function of the expectancies built up prior to that point. Such proactive 'thinking ahead' characterizes coreference models like the Expectancy Hypothesis (Arnold, 2001) as well as a more recent Bayesian approach (Kehler et al., 2008)—in both cases, properties of the discourse may directly or indirectly influence comprehenders' expectations about subsequent mention of a referent. This would be consistent with evidence of prediction in language processing that has accrued at other levels of linguistic representation, including morphophonology, syntax, and semantics (e.g., Federmeier, 2007; Kuperberg & Jaeger, 2016).<sup>2</sup> For example, when hearing *The boy will eat the...*, listeners look more at edible than inedible objects in a visual array, and this preference emerges before they hear the noun (...*cake*) (Altmann & Kamide, 1999). There is good

---

<sup>2</sup> While the extent of the role that prediction plays in language processing is under current debate (Huettig & Mani, 2016; Nieuwland et al., 2017), the contribution of a proactive component to human cognition is widely agreed upon (Bar, 2007).

evidence that listeners look at referents when they are named, or when there is syntactic and/or semantic information signaling that they are about to be named (Kamide, Scheepers, & Altman, 2003; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In all of these cases, the relevant cue (e.g., *eat*) can be pinpointed in the speech signal, and looking behavior predicted and examined in well-defined timeframes immediately following that cue.

Pinpointing the cue that focuses attention on referents in a mental model is less straightforward. Presumably the construction of event representations emerges from a complex combination of cues. This makes it challenging to detect effects of these models on gaze allocation because of their potential to be more widely distributed over time. Indeed, visual world eye-tracking experiments that test for predictive looks contingent on IC bias have produced inconsistent results (Cozijn, Commandeur, Vonk, & Noordman, 2011; Itzhak & Baum, 2015; Pykkönen & Järvikivi, 2010). Our study targets anticipatory coreference processing, but we use a cue whose effects emerge in a different way from those in IC studies. In IC experiments, the reported bias to a causally-implicated referent is taken to reflect the assignment of thematic roles. We hold constant thematic role, and indeed all referent properties typically enumerated in the pronoun interpretation literature, while manipulating grammatical aspect on the verb. The assignment of Source and Goal roles does not vary, nor do their grammatical roles, their relative recency, their parallelism with a subsequent subject pronoun, etc. We manipulate nothing about the pragmatic status or morphosyntactic encoding of the referents at all (i.e., the types of referent properties typically implicated in next-mention expectancies and assessed in corpus studies; Arnold 1998); rather aspectual effects arise via a

manipulation of the completed/ongoing nature of the event and the repercussions that has on the discourse and the comprehenders' model thereof. Note that an effect of aspect need not arise in all contexts. It is only when the status of an event is ongoing that it has the capacity to minimize the relevance of a protagonist associated with the end state. A coreference model for capturing such behavior would therefore depend on conditional activation of certain factors, particularly those for which the causal relationship with pronoun interpretation is indirect (see Kehler & Rohde, 2017). This makes comprehenders' potential use of this cue all the more impressive for real-time computation regarding upcoming next mention. Crucially our study targets the *anticipatory* use of aspect, extending beyond the integration effects reported in prior coreference work on aspect. To test for a potentially broadly distributed effect of aspect on listeners' next-mention expectations in transfer-of-possession contexts, we measure looks to event participants during an intersentential pause before the onset of a subsequent sentence.

## **2. Method**

### *2.1. Participants*

Sixty-three University of Hawai'i students who identified as native speakers of English participated after giving informed consent. Data from 7 participants was excluded prior to analysis due to eyetracker calibration difficulty ( $n=4$ ), non-normal vision or hearing ( $n=2$ ), or noise interference ( $n=1$ ). Data from 3 participants was excluded after data inspection, due to insufficient data points in the eye gaze record (see

2.3), leaving 53 participants (28 females, mean age 23) in the final analysis. The study was approved by the UH Human Studies Program.

## 2.2. *Materials and procedure*

Linguistic stimuli consisted of two-sentence discourses followed by a question, as in (1). In the 20 experimental items, context sentences contained a transfer-of-possession verb in a double-object construction. A 2×2 Latin Square design manipulated grammatical *Aspect* in the context sentence (perfective/imperfective) and *Reference* of the subject pronoun in the continuation (Source/Goal of context sentence, disambiguated through gender), with 5 items per condition distributed across 4 lists, each presented in two orders. Six practice items and 40 fillers were analogous to the experimental items in their overall structure (context-continuation-question) but described non-transfer-of-possession events (half perfective, half imperfective). Continuations in fillers started with pronouns referring to human or non-human participants in the context sentence. In experimental items, the final question asked about the referent of the gender-disambiguating pronoun in the continuation. Questions in fillers asked about different aspects of the context or continuation sentences. (See Appendix A for materials.) All materials were recorded by a female native speaker of American English using a clear speaking style.<sup>3</sup> Visual scenes contained 3 areas of interest (AOIs), representing the

---

<sup>3</sup> The pronunciation of the auditory materials and timing of all stimuli (see below) were chosen to be appropriate for second-language learners of English, who participated in a related experiment, while still natural for native speakers of English. Sample recordings are available at: <http://www2.hawaii.edu/~aschafer/snds.html#GTRS>



Source, Goal and Theme (Figure 1, panel A). Location of AOIs was counterbalanced across items.

- (1) Donald {brought/was bringing} Melissa a fancy drink. [context sentence]  
{He/She} obviously liked hosting parties. [continuation]  
Who liked hosting parties?

[INSERT FIGURE 1 ABOUT HERE]

The experiment utilized an SMI RED250 eye-tracker sampling at 250Hz. Each trial began with a 2,000 ms display of the visual scene, followed by the context sentence ( $M = 2,856$  ms,  $SD = 277$ ), a 2,500-ms intersentential pause, the continuation sentence ( $M = 2,755$ ms,  $SD = 331$ ), a 1,750-ms pause, and the question. 250 ms into the second pause, the AOIs were replaced with empty boxes (Figure 1, panel B). Participants were instructed to click on the box corresponding to the position of the event participant that best answered the question. This memory component was added to engage participants and disguise the objective of the experiment.

### 2.3. *Data treatment and analysis*

Eye gaze and mouse-click responses were recorded and exported through SMI *Experiment Suite* software. Mouse-click accuracy in experimental items was close to

ceiling (98.4%) and thus not further analysed. Eye gaze data were classified as fixations, saccades and blinks using the software's default settings. Data were binned into 20-ms samples for further analysis. Trials with a very low proportion of fixations overall were removed (4.2% of data).<sup>4</sup> Participants with fewer than 15 (out of 20) experimental trials remaining after this procedure were excluded from further analysis ( $n=2$ ). One additional participant was excluded for consistently fixating the Theme during the intersentential pause, thus rendering Source/Goal analysis impossible.

Our analysis tests whether grammatical aspect in the context sentence modulates listeners' looks to the Source versus the Goal of the transfer-of-possession event *before* they hear a pronoun. On the standard assumption that it takes approximately 200 ms to execute a ballistic eye movement in response to an acoustic stimulus (Matin, Shao, & Boff, 1993), our region of interest extends to 200 ms after the onset of the pronoun in the continuation. It begins 500 ms after the offset of the context sentence to allow for sentence wrap-up and for looks directly contingent on the naming of entities in that sentence to dissipate. In the absence of predictions for the effect of time within this region of interest ('Silence'), we calculated a 'GoalAdvantage' score for the entire region for each trial by subtracting the number of 20-ms bins with looks to Source from those with looks to Goal.<sup>5</sup> A GoalAdvantage score was also calculated for the region from 200

---

<sup>4</sup> To identify trials with insufficient fixation data, we calculated for each trial the proportion of sample points containing fixations out of all sample points. Trials with values more than 2 *SDs* below the mean were excluded.

<sup>5</sup> This dependent measure was chosen over weighted empirical logits (Barr, 2008) calculated at the trial level because it and the model residuals were more normally

to 1,500 ms after pronoun onset ('Continuation') in order to examine the effect of the disambiguating pronoun and potential interactions with aspect. We used linear mixed-effect regression with maximal random effects structures (Barr, Levy, Scheepers, & Tily, 2013) to model the effects of *Aspect* (perfective/imperfective; contrast-coded), *Reference* (pronoun disambiguating to Source/Goal; contrast-coded), and *Window* (Silence/Continuation; treatment-coded, reference level = Silence) on GoalAdvantage.

### 3. Results

Figure 2 shows an overview of participants' fixations over the entire trial, collapsing over experimental manipulations. Visual inspection indicates a small but sustained bias to look at the Goal versus Source during the intersentential pause (and beyond), consistent with the known bias to remention Goals following transfer-of-possession events (Arnold, 2001; Stevenson et al., 1994).

[INSERT FIGURE 2 ABOUT HERE]

Of key interest here is whether this bias is modulated by grammatical aspect. Figure 3 depicts looks to Goal and Source for each Aspect-by-Reference condition. Visual inspection suggests a Goal-bias following perfective but not imperfective aspect, in both distributed. Models fit to empirical logit data aggregated by participant and by item also met model assumptions well, and produced similar patterns of significance.

Reference conditions. This is supported by the output of the statistical model (see Appendix B, Table B1). Critically, a significant effect of Aspect emerges during the Silence region ( $b = -5.4, t = -2.3, p = .02$ ), with fewer looks to Goal following imperfective than perfective aspect. The only other effect that reached significance was the interaction between Window and Reference, reflecting that, as expected, looks to Goal (vs Source) were modulated by the disambiguating pronoun at the onset of the Continuation ( $b = 8.6, t = 2.9, p = .004$ ).<sup>6</sup>

[INSERT FIGURE 3 ABOUT HERE]

#### 4. Discussion

This study used the visual world paradigm (VWP) to measure listeners' coreference expectations during an intersentential pause prior to the encounter of a pronoun in a subsequent sentence. Given that no single acoustic event determines precisely when listeners' mental event representations may determine coreference

---

<sup>6</sup> In order to probe potential effects of Aspect during the Continuation region, we ran the same model with the reference level for Window reset to 'Continuation'. In addition to the same Window-by-Reference interaction, this model returned a significant simple effect for Reference during the Continuation region, indicating more looks to Goal when pronoun gender disambiguated reference to the Goal. Neither the simple effect of Aspect, nor its interaction with Reference were significant (Appendix B, Table B2).

expectations, we analysed fixations aggregated over the largest meaningful time period for the hypothesis at stake. This 2,200-ms window shows a sustained preference for fixating the Goal over the Source, and this preference was greater following perfective than imperfective transfer-of-possession events. Crucially, it appeared well before the onset of the pronoun. The findings lend direct support to the claim that the focus of attention on particular event participants in comprehenders' mental models allows them to proactively create expectations about who or what is likely to be mentioned next in the upcoming discourse. We have shown that a subtle linguistic manipulation of the way an event is characterized can affect listeners' expectations about upcoming reference, that these expectations are not reliant on the encounter of a pronoun or cue regarding the upcoming sentence's role in the discourse (e.g., a connective like *because* for IC verbs), and that they can be captured with VWP methods despite the more tenuous nature of the link that we must assume between mental event representations and eye movements.

These findings align well with Van Berkum, Koornneef, Otten, and Nieuwland's (2007) 'immediate focusing' account of the IC phenomenon (see note 1), which holds that comprehenders "use the implicit causality cue in something like 'David praised Linda because...' *proactively*, and essentially *predict*, before the pronoun comes along, that the remainder of the sentence will tell us something about Linda" (p. 167, italics in the original). However, despite suggestive evidence from other experimental paradigms (ERP, eye-tracking while reading; Van Berkum et al., 2007; Koornneef & Van Berkum, 2006), VWP studies testing for proactive IC-driven predictions of upcoming discourse referents have been somewhat inconclusive (Cozijn et al., 2011; Itzhak & Baum, 2015; Pyykkönen & Järvikivi, 2010). While these studies found evidence for differential looks

to event participants conditioned on IC bias in time segments including or immediately following the causal connector and the pronoun, effects *preceding* the connector and the pronoun have only been reported in one study (Pyykkönen & Järvikivi, 2010), and this effect has been questioned as a potential experimental artifact since pronouns always referred to bias-consistent antecedents in that study (Cozijn et al., 2011). By contrast, our VWP study shows evidence of proactive prediction of upcoming referents in the context of transfer-of-possession events and the manipulation of the portrayal of the event as completed or ongoing, in an experimental design that systematically crossed reference of the pronoun with aspect to eliminate any reinforcement of bias. It is possible that the greater length of our silence region provided a sufficient window for such expectations to emerge before the arrival of disambiguating material.

This study shows that the visual world paradigm can be used to capture discourse-level effects of this nature, with choices in design and analysis that reflect the more tenuous linking assumptions that must be made between incrementally updated mental event representations during discourse processing and allocation of eye gaze in a visual world. Exploring a wider variety of discourse contexts using this paradigm should be fruitful in future work aimed at understanding the dynamic mutual relationships between mental models of events and processing of discourse in real time.

## **Acknowledgments**

This work was supported by the National Science Foundation (BCS-1251450 to T.G. & A.J.S.). We thank A. L. Blake, Amber Camp, Victoria Lee, Wenyi Ling, Ivana Matson,

Eric Stepan, and Maho Takahashi for assistance with various aspects of this experiment, and the reviewers and associate editor of this journal for their helpful feedback.

### **Supplementary material**

Appendix A: Linguistic materials.

Appendix B: Output of statistical models.

### **References**

- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264. doi: 10.1016/S0010-0277(99)00059-1.
- Arnold, J. E. (1998). *Reference form and discourse patterns*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, 31, 137–162. doi: 10.1207/S15326950DP3102\_02.
- Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280-289. doi: 10.1016/j.tics.2007.05.005

- Barr, D. J. (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59, 457-474. doi: 10.1016/j.jml.2007.09.002
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing. *Journal of Memory and Language*, 68, 255-278. doi: 10.1016/j.jml.2012.11.00
- Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language*, 64, 381–403. doi:10.1016/j.jml.2011.01.001
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44, 491–505. doi: 10.1111/j.1469-8986.2007.00531.x
- Ferretti, T. R., Rohde, H., Kehler, A., & Crutchley, M. (2009). Verb aspect, event structure, and coreferential processing. *Journal of Memory and Language*, 61, 191–205. doi: 10.1016/j.jml.2009.04.001
- Ferstl, E. C., Garnham, A., & Manouilidou, C. (2011). Implicit causality bias in English: a corpus of 300 verbs. *Behavior Research Methods*, 43, 124–135. doi: 10.3758/s13428-010-0023-2
- Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry*, 5, 459-464.
- Hartshorne, J. K. (2014). What is implicit causality? *Language, Cognition and Neuroscience*, 29, 804-824. doi: 10.1080/01690965.2013.796396



- Huetting, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, 31, 19-31. doi: 10.1080/23273798.2015.1072223
- Itzhak, I., & Baum, S. R. (2015). Misleading bias-driven expectations in referential processing and the facilitative role of contrastive accent. *Journal of Psycholinguistic Research*, 44, 623-650. doi: 10.1007/s10936-014-9306-6
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32, 37-55. doi: 10.1023/A:1021933015362
- Kehler, A., Kertz, L., Rohde, H., & Elman, J. (2008). Coherence and coreference revisited. *Journal of Semantics* 25, 1-44. doi: 10.1093/jos/ffm018
- Kehler, A., & Rohde, H. (2017). Evaluating an expectation-driven QUD model of discourse interpretation. *Discourse Processes*, 54, 219-238. doi: 10.1080/0163853X.2016.1169069
- Koornneef, A., Dotlačil, J., van den Broek, P., & Sanders, T. (2016). The influence of linguistic and cognitive factors on the time course of verb-based implicit causality. *The Quarterly Journal of Experimental Psychology*, 69, 455-481. doi: 10.1080/17470218.2015.1055282

- Koornneef, A. W., & Sanders, T. J. (2013). Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes*, 28, 1169–1206. doi: 10.1080/01690965.2012.699076
- Koornneef, A. W., & Van Berkum, J. J. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54, 445–465. doi: 10.1016/j.jml.2005.12.003
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31, 32–59. doi: 10.1080/23273798.2015.1102299
- Madden, C. J., & Ferretti, T. R. (2009). Verb aspect and the mental representation of situations. In Klein, W., & Li, P., (Eds.) *The expression of time* (pp. 217–240). Mouton de Gruyter. doi: 10.1515/9783110199031.217
- Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes*, 29, 83–112. doi: 10.1207/S15326950dp2902\_1
- Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception, & Psychophysics*, 53, 372–380.
- Moens, M., & Steedman, M. (1988) Temporal Ontology and Temporal Reference. *Computational Linguistics*, 14, 15–28.
- Nieuwland, M., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Zu Wolfsturn, S.V.G., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D., Rousselet, G., Ferguson, H., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., Donaldson, D., Kohút, Z., Rueschemeyer, S., & Huettig, F. (2017).

Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology. *bioRxiv*, 111807; doi: <https://doi.org/10.1101/111807>

Pyykkönen, P., & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psychology*, 57, 5–16. doi: 10.1027/1618-3169/a000002

Rohde, H., Kehler, A., & Elman, J. L. (2006). Event structure and discourse coherence biases in pronoun interpretation. In R. Sun (Ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 617-622). Mahwah, NJ: Lawrence Erlbaum Associates.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus, and the representation of events. *Language and Cognitive Processes*, 9, 519–548. doi: 10.1080/01690969408402130.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634. doi:10.1126/science.7777863

Van Berkum, J. J., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, 1146, 158–171. doi: 10.1016/j.brainres.2006.06.091

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

## Figure legends

*Figure 1.* Sample visual displays.

*Figure 2.* Overview of the time course of fixations across the trial, collapsed over Aspect and Reference. Proportions to each AOI are calculated out of all fixations to an AOI. Trials were aligned by the onset of the Continuation. (In half the items the subject of the Continuation refers to the Source, in the other half to the Goal. Hence when collapsing over Reference, roughly equal fixations to Source and Goal are expected in the Continuation.)

*Figure 3.* Proportion of looks to Source and Goal for each Aspect-by-Reference condition. Grey shading indicates region of interest (Silence). Error bands indicate 95% confidence intervals of the participant means adjusted for repeated measures.

Figure 1

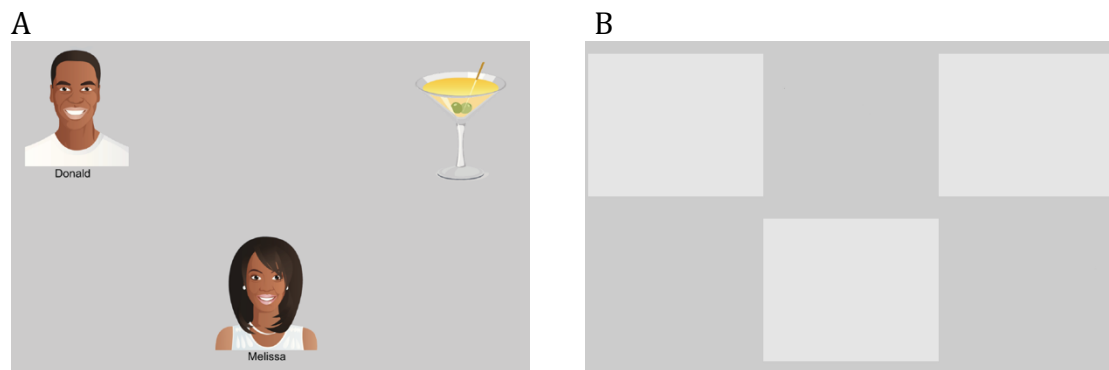


Figure 2

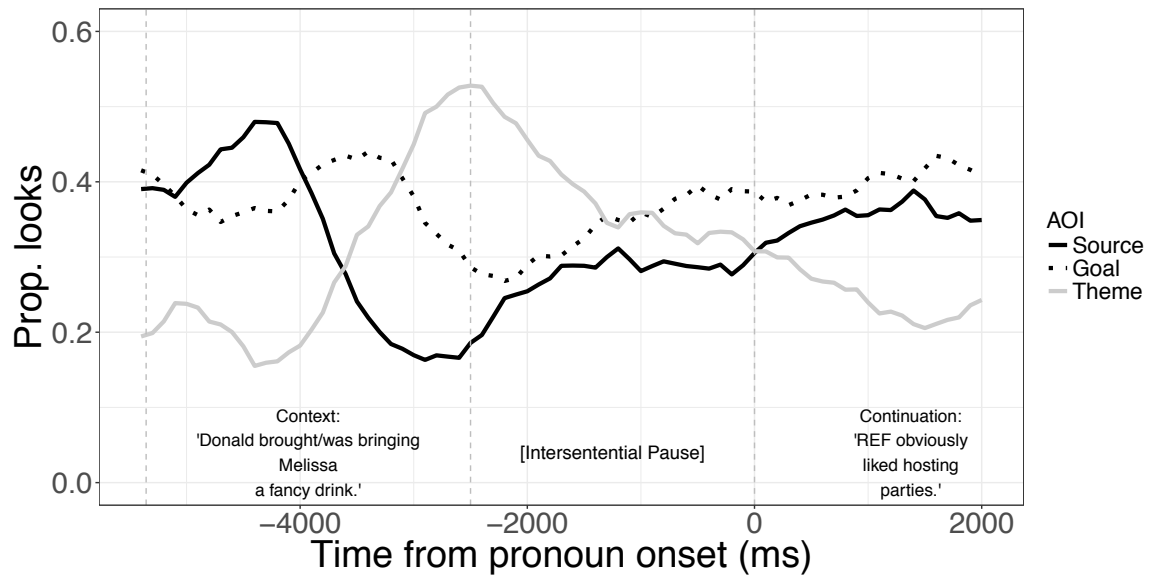


Figure 3

